

Міністерство освіти і науки України Національний технічний
університет України «Київський політехнічний інститут» Фізико-
технічний інститут



Комп'ютерний практикум №1

З дисципліни: "Криптографія"

Тема: "Експериментальна оцінка ентропії на символ джерела
відкритого тексту"

Перевірила:

Селюх К. І.

Виконали:

студенти III курсу

групи ФБ-95

Корольова В.Р.

групи ФБ-96

Гуменюк О.О.

Мета: Застосовувати практичні навички, при оцінюванні ентропії на символ джерела, з дослідженням ентропії на символ джерела та надлишковості, Досягти розуміння та правильного порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії.

Постановка задачі: У даному практичному практикумі, досліджуємо частотності букв та біграм російської мови, ентропії відповідно до букв(H_1) та біграм(H_2), щодо різних ситуацій по відношенню до тексту, тобто чи ми будемо обробляти з/без пробілами/ів з заміною ймовірностей на частотами.

Хід роботи:

- 1) Написали програму для підрахунку частот букв/біграм.
- 2) Підраховували H_1 та H_2
- 3) Підраховували частоти біграм та H_1 та H_2 відповідно для тексту з/без пробілами/ів.
- 4) Оцінили значення $H(10)$, $H(20)$, $H(30)$
- 4) Оцінили надлишковість російської мови в різних мовних джерелах

Труднощі_1:

Вивід файлу-некоректне перенесення даних до заданого масиву.

Рішення:

Ми повинні були зчитувати текстовий файл та переносити дані в масив. Після цього обробляли масив зручним чином.

Труднощі_2:

Проблема з переведення у нижній регістр букви верхнього регістру російської мови. Ми вирішили взяти функцію `tolower()`, проте вона спрацьовувала неадекватним чином на російських літерах.

Рішення:

Написали свою функцію, **`tolower_rus()`**

Що приймала значення, оброблювала та виводила коректно у інший файл. .

Труднощі_3:

Проблема з видалення лишніх символів.

Рішення:

Зробили масив з літер з циклом, при якому перевірялись дані нашого масиву і ,якщо елемент не є однією з літер масиву, то ми його видаляли . Тут було чимало складнощів ,тому прийняли рішення ,що треба зробити через булівську функцію, що приймає значення і порівнює символи ,якщо вона виводить значення **`true`**, то функція не видаляє літеру, якщо **`false`**, то відбувається видалення. Далі ми це ставимо у функцію вводу до файлу ,використовуємо ту ж саму функцію, яку використовували при видаленні регістрів. Записуємо значення масиву даних до файлу.

Труднощі_4:

Проблема з біграмами.

Складнощі з коректним записом даних ,адже було декілька варіантів аби записувати дані, як елементи масиву та робити це циклічно ,щоб воно записувало, проте виникла проблема з адекватним записом до файлу.

Рішення:

Функція для передачі виводу з консолі до файлу ,що мала назву **`freeopen()`**.

Труднощі_5:

Проблема:після того,як ми закінчували вивід значень з консолі до файлу,навіть ,коли сама функція скінчилась і потік її також закривався.У нас після цього,було не можливо вивести щось у консоль,воно не записувалось у файл, ні виводилось у консоль.

Рішення:

Використання буферу ,за допомогою функцій `ifstream/ofstreambuff` тим самим способом.

Таблиці частот букв

Букви без пробілів

а	88028	0,070931	-0,27077	Сумма	Ентропія	R
б	23104	0,018617	-0,10699	1,00000008	4,454569	0,100849
в	55357	0,044605	-0,20013			
г	21134	0,017029	-0,10006			
д	37445	0,030172	-0,15239			
е	111698	0,090004	-0,31266		+	
ж	12519	0,010088	-0,06689			
з	18595	0,014983	-0,09081			
и	91995	0,074127	-0,27826			
й	15666	0,012623	-0,07962			
к	34745	0,027997	-0,14442			
л	50544	0,040727	-0,18807			
м	40027	0,032253	-0,15979			
н	84597	0,068166	-0,26413			
о	142140	0,114533	-0,35805			
п	31149	0,025099	-0,13343			
р	55358	0,044606	-0,20013			
с	69033	0,055625	-0,23185			
т	85262	0,068702	-0,26543			
у	29874	0,024072	-0,12942			
ф	2015	0,001624	-0,01505			
х	11995	0,009665	-0,06469			
ц	6557	0,005283	-0,03997			
ч	19092	0,015384	-0,09265			
ш	9489	0,007646	-0,05376			
щ	4583	0,003693	-0,02984			
ы	24202	0,019501	-0,11077			
ь	25486	0,020536	-0,11512			
э	5527	0,004454	-0,03479			
ю	9151	0,007374	-0,05223			
я	24672	0,01988	-0,11237			

Букви з пробілами

	251837	0,168693	-0,43312	Сумма	Ентропія	R
а	88028	0,058965	-0,24081	1,000001	4,357822	0,120378
б	23104	0,015476	-0,09307			
в	55357	0,037081	-0,17625			
г	21134	0,014157	-0,08696			
д	37445	0,025083	-0,13337			
е	111698	0,074821	-0,27986			
ж	12519	0,008386	-0,05784			
з	18595	0,012456	-0,07881			
и	91995	0,061623	-0,24775			
й	15666	0,010494	-0,06899			
к	34745	0,023274	-0,12626			
л	50544	0,033857	-0,16537			
м	40027	0,026812	-0,13998			
н	84597	0,056667	-0,23468			
о	142140	0,095212	-0,32303			
п	31149	0,020865	-0,11648			
р	55358	0,037081	-0,17625			
с	69033	0,046242	-0,20507			
т	85262	0,057113	-0,23588			
у	29874	0,020011	-0,11292			
ф	2015	0,00135	-0,01287			
х	11995	0,008035	-0,05592			
ц	6557	0,004392	-0,03439			
ч	19092	0,012789	-0,08043			
ш	9489	0,006356	-0,04639			
щ	4583	0,00307	-0,02563			
ы	24202	0,016212	-0,09641			
ь	25486	0,017072	-0,10025			
э	5527	0,003702	-0,0299			
ю	9151	0,00613	-0,04505			
я	24672	0,016527	-0,09782			

Таблиці частот біграм

Мал.Біграми без пробілів з кроком_1(перетинається)

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я
а	220	2158	5934	1499	2692	2589	1608	4656	1437	538	6501	7692	4932	9374	901	2221	5949	6792	8049	517	220	1108	1567	1859	2195	421	0	0	527	1380	2492
б	691	6	43	2	11	2932	14	4	629	0	61	1266	66	469	5064	6	1808	771	3	1833	1	290	1	2	16	705	5565	321	44	22	458
в	6702	362	589	427	566	7212	94	481	4673	0	948	2122	443	2921	11484	1117	2494	5350	1436	1016	40	80	88	177	545	15	2738	176	655	6	400
г	1742	48	74	41	1786	1299	1	11	1241	0	153	1528	49	343	10652	67	1277	128	24	527	9	4	0	76	15	0	0	0	16	0	23
д	8136	56	1169	103	154	7187	28	44	3503	0	293	1460	98	2723	4868	361	1218	1089	327	2082	7	48	87	156	91	0	886	799	44	14	414
е	404	2588	5439	3989	5935	2756	1677	2360	2249	4809	2360	8722	8481	16898	2233	4381	8815	10178	8156	785	169	1227	976	2453	1213	880	0	0	524	275	766
ж	1071	53	6	5	1739	5514	6	1	1479	0	95	6	5	2099	38	4	9	12	1	291	1	0	2	13	0	0	0	67	0	0	2
з	5230	422	1149	260	1073	550	74	125	666	0	229	398	1138	2515	1139	148	683	217	125	946	8	7	45	45	36	0	773	122	110	1	361
и	1069	2339	5658	1303	2856	5662	838	4412	4572	1979	4137	4172	6123	6865	2564	2959	2189	6067	9255	518	238	4190	813	3509	770	196	0	0	589	1146	5007
й	250	478	914	388	687	165	249	250	947	0	707	193	673	1824	695	1490	552	2553	732	220	141	114	239	423	471	2	0	0	195	12	102
к	6500	207	402	91	120	399	215	62	4505	0	227	745	219	812	13302	275	2105	758	1601	1675	22	23	77	171	23	0	0	0	173	1	35
л	5278	158	284	165	294	6147	910	50	10618	0	445	308	81	702	7038	224	52	950	209	1871	13	12	6	170	10	1	520	9465	39	1721	2803
м	3984	569	964	343	557	6245	96	240	5259	0	651	448	626	2371	6174	1306	300	1565	517	3056	80	77	227	781	53	1	2565	38	188	78	668
н	15846	157	155	365	544	12551	15	56	14981	0	326	17	42	6326	16836	180	89	1344	1836	2258	31	85	920	304	12	42	6746	979	25	142	1387
о	208	9021	13166	8219	9026	4733	3139	2906	3055	6630	3468	10524	8095	9266	2686	5292	8604	13401	10544	827	478	805	591	3476	844	219	0	0	763	902	1252
п	2229	8	9	0	2	3030	0	2	895	0	57	780	4	288	11154	129	10904	54	59	874	5	0	10	12	2	0	415	17	2	0	208
р	10994	106	668	804	218	9862	623	24	5888	0	474	255	1358	945	11212	242	161	1522	1110	3247	16	256	188	118	569	4	1766	1336	12	44	1336
с	2494	182	2526	94	408	4860	73	31	2422	0	5029	3897	874	1616	6449	2522	685	1703	20910	2529	87	260	112	691	165	0	638	2193	132	207	5244
т	8616	346	7731	166	923	8438	94	156	9276	0	961	346	267	2537	21299	1037	3633	4030	620	2288	70	41	77	498	16	17	1315	9485	217	17	745
у	384	800	1071	1359	3997	453	2126	382	484	35	1080	998	1369	1331	372	1771	1177	2063	2251	135	44	306	44	2084	436	883	0	0	152	2184	103
ф	323	0	3	4	1	271	0	0	340	0	1	75	2	3	303	1	465	118	9	41	12	0	0	2	0	0	39	0	1	1	0
х	461	222	725	407	362	199	82	199	749	0	394	352	310	808	2695	1052	576	1113	319	390	99	20	69	144	53	0	0	0	138	9	48
ц	5007	7	101	18	14	1625	2	5	2963	0	522	11	8	28	257	43	5	18	16	153	3	0	0	6	0	0	236	0	6	0	3
ч	3000	1	42	1	5	6013	3	2	2756	0	259	117	8	1357	28	6	112	12	4387	575	1	0	0	9	269	0	0	128	1	0	0
ш	821	12	24	6	4	3752	1	1	2442	0	266	426	13	659	280	26	4	24	138	231	1	0	1	3	2	0	0	329	1	21	1
щ	419	0	0	0	0	2788	0	0	1117	0	0	0	1	112	0	2	4	2	0	107	0	0	0	0	0	0	0	0	31	0	0
ы	89	457	2077	216	611	2642	57	271	664	1542	577	2430	2500	1427	493	987	427	1700	1486	229	30	2444	24	253	270	7	0	0	189	3	100
ь	120	1213	1459	257	535	870	87	582	924	0	3091	154	597	5133	767	1091	275	3142	1376	274	56	90	70	600	1360	10	0	0	316	617	420
э	0	4	2	36	2	0	0	1	0	5	140	98	8	74	0	122	26	18	4983	0	7	1	0	0	0	0	0	0	0	0	0
ю	57	413	405	139	1123	81	113	523	289	0	256	81	197	459	292	507	305	519	1581	72	51	13	203	430	14	830	0	0	96	59	43
я	182	711	2568	427	1200	873	294	758	974	128	1037	923	1440	2312	865	1580	455	1820	3202	307	75	494	120	627	39	350	0	0	373	289	251

Мал.Біграми без пробілів з кроком_2(не перетинається)

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я	
а	105	1071	2987	736	1338	1250	771	2320	723	277	3278	3848	2469	4711	451	1071	2953	3466	4016	261	101	525	783	904	1112	220	0	0	247	692	1259	
б	353	3	20	2	3	1536	7	2	310	0	33	624	32	220	2558	1	913	383	0	913	1	140	0	0	7	353	2753	152	17	11	221	
в	3283	177	296	223	308	3628	38	239	2294	0	511	1058	229	1479	5749	558	1241	2687	696	526	20	38	35	77	258	5	1402	91	333	2	183	
г	870	31	39	16	902	663	0	5	613	0	73	768	26	172	5352	35	617	64	10	274	6	4	0	36	6	0	0	0	9	0	10	
д	4087	20	570	50	76	3555	17	14	1749	0	137	710	55	1387	2464	177	618	551	172	1016	3	20	44	77	49	0	421	388	26	6	205	
е	199	1272	2705	1993	2969	1348	829	1141	1102	2414	1169	4437	4271	8319	1138	2191	4426	5029	4143	401	77	625	497	1201	609	431	0	0	257	138	380	
ж	574	24	3	2	875	2840	4	0	732	0	49	5	1	1079	18	3	5	7	1	149	0	0	1	7	0	0	0	37	0	0	1	
з	2624	193	579	128	546	265	38	65	338	0	119	214	561	1264	578	68	336	110	65	466	4	4	23	19	16	0	380	59	57	0	173	
и	508	1151	2833	651	1431	2853	426	2254	2284	1000	2052	2085	3020	3417	1296	1498	1079	3007	4691	270	112	2114	419	1720	360	104	0	0	292	577	2467	
й	122	237	459	188	348	77	120	117	479	0	378	88	315	889	358	674	284	1306	372	112	66	54	130	217	231	2	0	0	95	7	56	
к	3286	106	200	44	50	214	106	31	2272	0	127	378	103	398	6614	142	1031	377	803	803	10	11	40	76	12	0	0	0	90	1	14	
л	2649	89	134	85	137	3095	449	21	5300	0	234	142	39	362	3476	110	28	477	97	951	5	8	5	81	5	1	279	4685	14	851	1407	
м	1991	286	497	186	281	3046	51	116	2652	0	315	220	314	1202	3097	658	156	809	269	1471	40	39	105	414	25	0	1309	15	99	40	343	
н	7961	79	83	199	268	6294	6	22	7496	0	177	9	13	3173	8419	99	44	700	913	1098	21	43	452	152	3	20	3357	492	12	69	692	
о	90	4522	6607	4082	4535	2374	1567	1467	1450	3371	1723	5265	4009	4636	1362	2590	4296	6681	5274	407	253	421	292	1778	403	115	0	0	387	443	613	
п	1150	4	4	0	0	1554	0	1	479	0	28	400	1	146	5566	68	5567	24	29	408	2	0	5	8	0	0	0	209	7	1	0	102
р	5535	64	344	404	111	4925	279	9	2913	0	225	126	680	470	5550	120	74	750	575	1591	7	123	98	67	287	2	865	685	7	21	642	
с	1207	109	1258	53	212	2406	31	15	1217	0	2459	1902	450	827	3270	1258	346	857	10484	1257	42	125	58	359	64	0	329	1077	63	105	2545	
т	4251	184	3837	72	473	4218	44	80	4665	0	465	174	126	1241	10595	514	1828	1982	312	1143	34	19	38	256	4	7	667	4674	111	7	373	
у	199	413	537	669	1988	230	1029	191	242	15	528	523	720	642	195	897	632	1050	1158	67	23	145	20	1048	218	456	0	84	1120	50	50	
ф	183	0	1	1	1	136	0	0	165	0	1	44	0	1	168	1	240	54	5	19	4	0	0	2	0	0	14	0	1	0	0	
х	226	118	356	208	172	104	35	102	371	0	204	180	161	410	1354	516	298	555	166	198	43	4	38	74	29	0	0	68	7	26	26	
ц	250	3	43	10	7	785	1	2	1517	0	260	7	4	11	128	14	2	11	8	71	2	0	0	3	0	0	127	0	3	0	1	
ч	1507	0	20	0	2	3038	2	0	1402	0	126	60	3	680	13	2	57	5	2136	292	1	0	0	3	139	0	0	64	0	0	0	
ш	457	7	15	2	3	1855	1	1	1223	0	136	206	7	325	138	12	2	7	84	112	0	0	1	0	2	0	0	176	0	10	0	
щ	205	0	0	0	0	1396	0	0	557	0	0	0	1	48	0	1	1	1	0	61	0	0	0	0	0	0	0	17	0	0	0	
ы	42	218	1039	106	317	1363	26	130	333	741	295	1234	1238	698	249	482	211	857	740	113	14	1200	9	138	136	4	0	0	112	2	43	
ь	54	573	733	122	282	440	35	309	498	0	1572	79	304	2562	382	546	137	1621	720	145	20	47	30	289	703	6	0	0	160	299	199	
э	0	2	1	17	1	0	0	1	0	0	3	68	48	6	40	0	65	7	12	2483	0	3	1	0	0	0	0	0	0	0	0	0
ю	28	228	176	68	549	43	51	264	157	0	129	43	88	236	154	250	158	260	801	29	23	10	100	209	6	405	0	0	47	27	25	
я	87	352	1317	216	596	456	139	384	491	64	535	451	735	1186	435	765	222	948	1645	161	37	252	64	325	23	164	0	0	178	152	131	

Мал. Біграми з пробілами з кроком_1(перетинається)

	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п	р	с	т	у	ф	х	ц	ч	ш	щ	ы	ь	э	ю	я		
	37087	2525	9611	20662	5569	9723	3873	1986	4733	13547	3	10194	3130	8337	22913	12905	21544	6174	20823	12972	4692	1354	1229	1017	6933	643	9	0	0	5385	97	2166	
а	16033	0	1396	4309	1136	1945	2226	1518	4235	283	535	5677	7489	4330	7627	71	711	5637	5267	7041	196	51	1038	1521	1535	2152	421	0	0	4	1370	2274	
б	183	683	3	32	0	10	2926	14	1	605	0	60	1265	62	463	5033	0	1803	765	0	1818	0	290	1	0	16	705	5565	321	0	22	458	
в	9919	6502	2	68	19	131	7046	0	344	4237	0	319	1975	144	2065	10888	343	2056	4295	255	922	0	22	13	12	504	14	2738	176	0	1	347	
г	725	1727	1	5	25	1737	1293	0	0	1208	0	109	1516	14	264	10617	0	1266	45	3	501	1	0	0	67	10	0	0	0	0	0	0	
д	1543	8114	8	976	66	95	7161	25	6	3405	0	211	1439	51	2476	4797	248	1177	952	227	2027	0	45	85	124	87	0	886	799	0	14	401	
е	26754	209	1399	2466	3172	4698	2331	1443	1595	671	4809	1244	8371	7132	14381	569	1477	8068	7560	6845	106	40	991	824	1788	1143	879	0	0	1	262	470	
ж	76	1069	47	0	2	1739	5511	6	0	1476	0	87	4	3	2085	32	0	7	4	0	290	0	0	2	12	0	0	0	67	0	0	0	
з	1730	5209	386	978	206	1008	530	70	66	598	0	120	373	1081	2288	1049	0	628	38	4	916	0	0	35	5	31	0	773	122	0	0	351	
и	25701	759	990	2968	710	1632	5177	688	3799	2958	1979	2849	3763	5199	4071	1119	306	1510	3678	7594	20	111	4060	687	2891	697	196	0	0	1137	4746		
й	12379	0	3	0	0	134	3	0	5	0	0	62	1	56	788	10	3	1	1356	196	0	0	42	84	130	412	1	0	0	0	0	0	
к	4356	6458	0	81	0	0	343	142	14	4194	0	17	694	6	248	13083	0	1880	339	1253	1580	2	0	39	2	14	0	0	0	0	0	0	
л	2577	5259	8	2	121	132	6086	904	6	10481	0	335	288	4	495	6933	36	0	486	59	1809	1	0	0	59	5	1	520	9465	1	1721	2750	
м	12130	3849	14	39	5	0	6082	0	2	4457	0	60	278	152	1178	5216	189	12	246	0	2713	2	0	184	18	0	1	2565	38	0	76	521	
н	1659	15826	17	9	338	467	12528	8	13	14838	0	263	0	0	6157	16749	1	46	1171	1740	2211	24	75	915	255	5	42	6746	979	3	141	1371	
о	30730	30	7237	10056	7526	7665	4191	2805	2252	1365	6630	2286	10125	6914	5803	614	2259	7783	10690	8779	53	318	595	524	2017	752	216	0	0	109	891	925	
п	101	2221	0	0	0	0	3029	0	0	889	0	52	773	0	271	11154	122	10896	52	55	873	4	0	10	6	2	0	415	17	0	0	207	
р	679	10977	69	601	793	196	9852	620	9	5841	0	436	242	1327	869	11175	147	149	1467	1074	3236	10	254	183	111	566	4	1766	1336	0	44	1205	
с	2731	2413	97	2390	12	228	4804	1	1	2318	0	4882	3854	782	1302	6254	2294	604	1519	20607	2475	51	247	94	648	159	0	638	2193	1	204	5227	
т	8877	8548	27	6286	7	430	8134	5	3	8692	0	540	210	43	1612	20769	67	3416	3149	52	2109	11	4	94	154	3	16	1315	9485	0	15	694	
у	6744	234	635	540	1193	3716	323	2020	219	12	35	758	917	1067	388	14	1118	993	1558	1879	17	6	265	26	1665	423	883	0	0	10	2180	36	
ф	16	323	0	0	0	0	271	0	0	340	0	0	75	1	3	303	0	465	117	7	41	12	0	0	1	0	0	0	39	0	0	1	0
х	8011	355	1	148	2	0	121	0	0	291	0	0	61	1	106	2241	1	297	131	11	215	0	0	0	0	2	0	0	0	0	0	0	0
ц	289	499	1	75	7	1	1618	0	0	2933	0	508	8	1	0	235	0	0	0	0	146	0	0	0	0	0	0	0	236	0	0	0	0
ч	129	2999	0	36	0	0	6012	0	0	2746	0	253	104	0	1332	21	0	110	0	4384	571	0	0	0	0	0	267	0	0	128	0	0	0
ш	211	819	0	2	0	0	3751	0	0	2433	0	262	424	1	588	270	4	2	10	134	229	0	0	0	0	0	0	0	0	329	0	20	0
щ	8	419	0	0	0	0	2788	0	0	1115	0	0	0	0	0	111	0	0	4	0	0	107	0	0	0	0	0	0	0	31	0	0	0
ы	8305	0	88	1231	67	162	2473	18	85	9	1542	222	2340	2231	373	0	119	229	985	955	0	0	2403	7	89	257	6	0	0	0	0	0	6
ь	12743	0	758	9	35	7	568	3	345	40	0	2502	1	133	3640	19	2	0	1789	567	0	10	0	36	49	1333	10	0	0	0	612	275	0
э	2	0	4	0	36	2	0	0	1	0	5	140	98	8	74	0	122	26	18	4983	0	7	1	0	0	0	0	0	0	0	0	0	0
ю	4426	1	229	1	19	959	13	41	423	0	0	10	13	6	118	0	0	58	82	1395	0	0	3	170	295	1	830	0	0	13	55	0	
я	14983	0	73	847	78	628	634	199	438	13	128	287	713	941	508	0	36	61	441	2191	1	0	426	56	224	7	349	0	0	0	288	122	

Одержані значення Н1 та Н2

	H	R
Букви з пробелами	4,357822	0,12037761
Букви без пробілів	4,454568806	0,100849355
Біграми з кроком 1(без пробілів)	4,106424237	0,171122019
Біграми з кроком 1(з пробілами)	3,952237358	0,209552528
Біграми з кроком 2(без пробілів)	4,106160822	0,171175189
Біграми з кроком 2(з пробілами)	3,951740412	0,202344807

$$H_{\infty} = \lim_{n \rightarrow \infty} H_n, \text{ где } H_n = \frac{1}{n} H(x_1, x_2, \dots, x_n),$$

а $H(x_1, x_2, \dots, x_n)$, в свою чергу, – ентропія n -грами відкритого тексту (x_1, x_2, \dots, x_n) :

$$H(x_1, x_2, \dots, x_n) = - \sum_{z_1, z_2, \dots, z_n} P(x_1 = z_1, \dots, x_n = z_n) \cdot \log_2 P(x_1 = z_1, \dots, x_n = z_n).$$

H10

Лабораторная работа №1

Произвольная часть текста:
взгляд

Использованные буквы:

Порядок n-граммы:
5 символов
1 символ
15 символов
20 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ: _____
Символ по счету: _____
Номер эксперимента: 51

Поле ввода символов:

Продолжить Другой

Неравенство для энтропии:
1.45702296890873 < H < 2.18752902292088

Двоичная таблица угаданных символов:
10000000000000000000000000000000 ^
100000000000000000000000000000000
000100000000000000000000000000000
000000000010000000000000000000000
10000000000000000000000000000000 v

Вероятности:

q[1]	= 0.6
q[2]	= 0.08
q[3]	= 0.1
q[4]	= 0.04
q[5]	= 0
q[6]	= 0
q[7]	= 0.04
q[8]	= 0
q[9]	= 0.02
q[10]	= 0
q[11]	= 0.02
q[12]	= 0.04
q[13]	= 0
q[14]	= 0
q[15]	= 0
q[16]	= 0
q[17]	= 0.02
q[18]	= 0
q[19]	= 0.02
q[20]	= 0
q[21]	= 0
q[22]	= 0.02
q[23]	= 0
q[24]	= 0
q[25]	= 0
q[26]	= 0
q[27]	= 0
q[28]	= 0
q[29]	= 0
q[30]	= 0
q[31]	= 0
q[32]	= 0

Строка состояния:

$$R0,708595 < R10 < 0,562494$$

H20

Лабораторная работа №1

Произвольная часть текста:
а_себя_труд_сравнил

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
25 символов
30 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ:

Символ по счету:

Номер эксперимента: 51

Неравенство для энтропии:
 $1.3025479110001 < H < 1.90703197875806$

Двоичная таблица угаданных символов:

Поле ввода символов:

Продолжить Другой

Вероятности:

q[1] = 0.62
q[2] = 0.14
q[3] = 0.1
q[4] = 0
q[5] = 0
q[6] = 0
q[7] = 0
q[8] = 0
q[9] = 0
q[10] = 0
q[11] = 0.02
q[12] = 0.02
q[13] = 0
q[14] = 0.04
q[15] = 0
q[16] = 0
q[17] = 0
q[18] = 0.02
q[19] = 0.02
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0.02
q[24] = 0
q[25] = 0
q[26] = 0
q[27] = 0
q[28] = 0
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0

Строка состояния:

$0,739490 < R_{20} < 0,618594$

H30

Лабораторная работа №1

Произвольная часть текста:
и_всегда_были_согласны_в_том_

Использованные буквы:

Порядок n-граммы:
5 символов
10 символов
15 символов
20 символов
25 символов
35 символов
40 символов
45 символов
50 символов

Введенный символ:

Символ по счету:

Номер эксперимента: 51

Неравенство для энтропии:
 $1.88667681617376 < H < 2.7288984057993$

Двоичная таблица угаданных символов:

Поле ввода символов:

Продолжить Другой

Вероятности:

q[1] = 0.5
q[2] = 0.14
q[3] = 0.08
q[4] = 0.04
q[5] = 0
q[6] = 0.02
q[7] = 0
q[8] = 0
q[9] = 0.02
q[10] = 0.02
q[11] = 0.02
q[12] = 0.02
q[13] = 0.02
q[14] = 0
q[15] = 0
q[16] = 0.02
q[17] = 0.02
q[18] = 0
q[19] = 0.02
q[20] = 0
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0.02
q[25] = 0
q[26] = 0.02
q[27] = 0.02
q[28] = 0
q[29] = 0
q[30] = 0
q[31] = 0
q[32] = 0

Строка состояния:

$0,622664637 < R_{30} < 0,454220319$

Оцінку надлишковості R російської мови у різних моделях відкритого тексту;

Висновки: Було засвоєнно навички з підрахунком ентропії на символ джерела, що дозволило розумітись на невизначеності інформації у відповідному тексті, досліджено ентропію, значення якої $=0$, у випадку, коли ми знали б яка була наступна буква в біграмі, відповідно при ентропії $=1$, ми не можемо знати результат. А також практично дослідили та проаналізували значення надлишковості російської мови, відповідно при різних конфігураціях.